Dialect Variation in Speaking Rate

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for graduation

with research distinction in Speech and Hearing Sciences in the undergraduate colleges of

The Ohio State University

By

Caitlin G. O'Neill

The Ohio State University

June 2008

Project Advisors: Dr. Robert Fox and Dr. Ewa Jacewicz

Abstract

The difference in speaking rates among American English regional dialects have been assumed and become popular belief in U.S. culture without supporting evidence to prove or disprove it. This study compares the speaking rates of those in south-central Wisconsin and western North Carolina in order to see if southerners do, in fact, speak more slowly than northerners. Age and gender are also compared. Secondly, the effect of dialect, age, and gender were examined in the difference in speaking rates in read sentences and spontaneous speech, because previous research has shown that speaking rate is slower when reading aloud than when having a conversation. Articulation rates were calculated and results showed a significant difference in both rates of spontaneous speech and read sentences for the two dialects, with Wisconsin being faster for both. Age was also a significant factor for both types of production in Wisconsin, and although gender did not have a main significant effect, there were significant differences in young men and young women in North Carolina for read sentences. These results supported some previous preliminary findings and should be considered in future dialect studies.

Table of Contents

Chapter 1

Introduction and Literature Review

1.1. Goals of the Study

The presence of many different versions or "dialects" of American English spoken across the United States is a well-known fact. These dialects can be defined as regional or social. Regional dialects refer to "varieties of a language that exist in different geographical areas" (Cran & NacNeil 2005). Social dialects are related to class, ethnicity, gender, or age (O'Neil 2006); an example of a social dialect is African American Vernacular English (Cran & MacNeil 2005). This study focuses on two regional dialects of American English, one spoken in south-central Wisconsin and the other in western North Carolina.

Linguists, while examining these regional dialects of American English, have found specific grammatical, phonological, and morphological differences among them. However, the fascination that Americans have with the lifestyle differences associated with these regions has lead to stereotypes about the dialects that are spoken and the people who speak them. One pervasive stereotype is that "southern speakers talk more slowly than northern speakers," (perhaps accompanying a stereotype that they think slower as well.) This may stem from the idea that southerners live their lives at a more relaxed pace than their northern neighbors. Despite the persistence of this idea of "slow-talking southerners" (Cran & MacNeil 2005) in our culture, there has been a lack of sufficient evidence provided by the academic community to support or disprove it. The same can be said for the stereotypes about age and gender, these stereotypes being that

older people speak more slowly than younger people, and women speak more quickly than men. While some research has been conducted to make some broad generalizations, there has not been specific instrumental evidence to support these familiar claims.

The current research uses data from eighty speakers, forty from south-central Wisconsin and forty from western North Carolina (each set from a three to four county range.) Each speaker group was composed of two age groups (18-34 years old and 51-65 years old) and equal numbers of male and female speakers. This large pool of speakers allows for a detailed and accurate examination of speaking rate to be measured, more so than any previous studies on speaking rate for these dialects. The composition of the speaker groups allows precise cross-dialectal, cross-generational, and cross-gender comparisons of speaking rate.

Another goal of the research was to determine the difference between the speaking rate and reading rate among these speakers, because previous research has shown there are significant differences. These comparisons will be made within-subjects and across dialects, by measuring the two sets of data each speaker provided for the study. This data consisted of an interview, in which the speaker spoke spontaneously about various topics, and a sentence reading set, in which the speaker read aloud two sets of sixty sentences provided for him.

1.2. Articulation Rate

In this study, "articulation rate" was actually measured rather than speaking rate. There is an important distinction between these two, although both are "defined as the number of output units per unit of time" (Tsao, Weismer, & Iqbal, 2006:1156). The

difference is that speaking rate includes pause intervals in its assessment and articulation rate does not. Therefore, articulation rate measures speech to determine length of actual sound production. In measuring the data in the present study, all pauses and hesitations were excluded by editing them out from the speakers' fluent phrases. The length of these phrases were measured and then divided by the number of syllables in the phrase in order to determine rate in syllables per second.

In an early study of articulation rate, Goldman-Eisler (1961) concluded that articulation rate was a "personality constant of remarkable invariance" (p. 173), meaning that both within and across individuals there was no significant differences in the rate of articulation. The overall rate was determined to be 3.7 words per second from a pool of speakers that were asked first to describe a cartoon and then to say what he or she thought the point of the story was. The speakers were then asked to repeat their explanation six more times. The description of the cartoon served as a "concrete situation" of speech, while the summary portion (the point of the story) was used as a process to have speakers speak spontaneously. A comparison of these two portions showed that there was virtually no difference between the "two levels of verbal operation" (p. 172) for any speaker as the articulation rate for the two portions of speech was almost exactly the same. The articulation rate did increase with repetition, leading Goldman-Eisler to remark that faster articulation rate could indicate speakers are using familiar, clichéd, or vernacular speech that is well practiced. However, these changes were not significant, especially compared to how much speaking rate changed (where pauses were included.) Therefore, it was concluded that the quantity and length of pauses were the major contributor to speaking rate.

More than twenty years later, a study by Miller, Grosjean, and Lomanto (1984) reevaluated the then well-established view that articulation rate was a minor factor in the reason for varying speaker rates, while pausing was the most influential aspect. By analyzing the interviews from thirty speakers, they found that "the average duration of a syllable can change substantially not only over the course of an entire interview or even a single response during an interview, but between adjacent stretches of pause-free speech" (p. 221). Results showed that in individual speakers, average syllable duration between the shortest and longest runs of speech differed by over 100 ms. For many speakers this difference was even higher. This study proved that articulation should not be dismissed when considering speaking rate, because it is, in fact, an important factor.

1.3. Reading Rate

Speaking rate and reading rate were compared by Crystal & House (1982) who noted that speakers using "less formal production" (p. 706) – meaning informal conversations rather than reading - showed more reduction in their speech, which increased their speaking rates. Hirose & Kawanami (2002) examined dialogue versus read speech, and focused on the prosodic differences that exist between the two, noting that "dialogue speech generally shows wider dynamic ranges in its prosodic features" (p. 97), such as in tone and rhythm, as well as a higher speech rate than read speech. Howell & Kadi-Hanifi (1991) also examined prosodic differences in reading and speaking. They had six speakers speak spontaneously (they were asked to describe a room of their choice), and then three months later, after transcribing their response, had them read what they had responded, as well as the response of two other speakers, in order to compare

speaking styles. They found that speakers produced a "larger number of short tone units" (p. 166) while reading, indicating more fragmentation and therefore a more formal style of speech. Stress position was also a factor, finding that speakers tended change the stress on vowels when reading: they did not stress some vowels that were stressed in the spontaneous speech, but added stress onto other vowels that were unstressed when spoken spontaneously. It was concluded that this variation in stress placement did not produce significant differences in speaking rate, but did observe that speaking rate "changes with the mode of delivery" (p. 168). Based on the findings of these studies, it is expected that the present results will show a difference between the articulation rate in read sentences and spontaneous speech.

1.4. Just Noticeable Difference

If this current study finds a significant difference in the articulation rates of any of the two groups of speakers (dialect, age, gender), an intriguing and perhaps important question to ask is, is that difference considerable enough for listeners to perceive it? For example, if Wisconsin speakers are shown to speak more quickly than North Carolina speakers but the difference is too small to be perceptible by a listener, then the stereotype would still be proven false. Quené (2007) provides insight as to how to discern whether or not listeners will observe this difference, called the Just Noticeable Difference (JND). In his study, 24 listeners listened to stimulus materials that had been sped up or slowed down by several increments. The listener would be presented with the normal stimuli paired with a manipulated one, and had to choose if the rates differed. They listened to 640 pairs in the duration of the experiment. In a second experiment, the listeners were

only given the choice of whether the first or second fragment in the pair was faster, as to eliminate response bias that could have occurred in asking to decipher equality and difference. The results of the study indicated that the JND was about 5%, meaning that a 5% or more difference of rate in the speech fragments could be perceived by the listener. This finding will be of use in consideration in this study's results.

1.5. Effects of Gender, Dialect and Age on Speaking Rate

Whiteside's 1996 study looked the characteristics of read speech in relation to gender. The study included three women and three men speakers with a British General Northern accent and each speaker read five sentences ten times each. Based on previous evidence that "men and women adopt different speaking styles in conversational interactions" (p. 23), the study explored what those stylistic components would mean for temporal features of speech. Results showed that women speakers had longer mean sentence durations, and also had higher standard deviations. However there was considerable variation among the speakers. One of the women had a significantly higher articulation rate than the two other women and one of the men had a significantly lower articulation rate than the other two men. When pausing was accounted for, it was found that women paused more frequently than the men. Due to the small subject size and the interspeaker variations found, these results must be viewed with caution. Whiteside suggested that results were considered preliminary and that the study should followed up with more speakers and larger amounts of data collected from each speaker.

A study of speaking rate across regions of the United States was conducted by Ray & Zahn (1990). The 93 speakers that participated in the study were from the

following regions: the Pacific Northwest, the Southwest, and the Upper Midwest. The speech samples analyzed were taken from two settings - public speaking and group discussion - both in regional university settings. Speaking rate was measured in words per minute and gender, region, and contextual comparisons were made. The only significant results that were found were in the context category (public versus discussion), meaning that there was little variation shown between gender and region.

It is important to note, however, that the division of speakers was done entirely by the university they attended, and the location of these universities ranged from Utah, Oregon, and Washington in the Pacific Northwest category, Texas and Louisiana in the Southwest, and Wisconsin and Ohio for the Upper Midwest. With such large areas being represented by one category, there is a chance for variability within categories (for example, recent research conducted by Jacewicz, Salmons, & Fox (2007) showed that there is a tendency for vowels in the dialect of Central Ohio to be longer than those of South-Central Wisconsin.) Also, the data was gathered by professors who taped certain students in the classroom setting. There was no mention if these students were originally from the category they were included in. The present study overcomes this problem because it only used speakers that had grown up in a narrowly defined location.

Social variables have also been examined for their role in speaking rate. The first study to look at speaking rate as a function of living in an urban or rural setting was conducted in 1998 by Hewlett & Rendall. While the present study does not look at urban and rural areas, it may be significant that the speakers from Wisconsin lived in a more urban setting than those in North Carolina. Hewlett & Rendall (1998) studied the claim that urban dwellers speak more quickly than rural dwellers by comparing the speech rates

of 24 speakers from either Orkney (a rural area of England) or Edinburgh (urban area.) The speakers were asked to read a passage and then were interviewed for several minutes about their lives (interests, future plans, etc.), and from this data both speaking and articulation rate was calculated in syllables per second (counting only the speaking turns with ten or more syllables). They found little difference in speaking rate while the speakers were reading, but found that in conversation mode the Edinburgh group was slower than in their reading rate, but this difference was not significant. Articulation rate also showed no significant differences in reading rate of the two groups but found the Orkney group to have a noticeably faster speaking rate in conversation than the Edinburgh group, which was significant. Thus, their results failed to support the claim that there is a difference in rate between urban or rural speakers. They concluded that exploring speaking rate differences across dialects rather than lifestyle may present more remarkable findings.

Byrd (1994) studied reduction in terms of speaker sex and dialect, which is important to this study because reduction "includes different types of simplification which speakers regularly exhibit in pronunciation" (p. 41). Types of reduction can include "assimilation, vowel centralization, and the deletion or simplification of segments" (p. 41). These simplifications can have an effect on speaking rate. For example, casual speech may be reduced, while formal speech may be slower because the speaker is over-articulating each word. Byrd explored how speakers of a certain sex and/or dialect use reduction compared to other groups, and how this affects speaking rate. Using the TIMIT database, a total of 2342 sentences read by 630 speakers were examined. Eight dialectal regions were found among them and they were separated

accordingly. Speaking rate was determined from the sentences and measured in syllables per second. Results showed that there was a significant difference between men and women (with men having a 6.2% faster rate), and also a significant effect of region, with the South region having the slowest speaking rate and the "Army Brat" category having the fastest. In terms of pausing, it was found that the Southern and South Midland speakers produced the most pauses and the North Midland, Western, and "Army Brats" produced less than was expected given a random distribution determined by a chi-square test. Men and women had no significant differences in their pauses.

Byrd recognized that there were some inadequacies in the speakers, due to the fact that the TIMIT database was 70% male and 30% female, making sex and dialect comparisons hard to make. Also, the divisions of the dialectal regions were too broad, leading Byrd to believe that the results in regards to dialect were inconclusive. Lastly, the speaking rates in this study were obtained from analyzing read sentences, not conversational data, which may have had an effect.

Verhoeven, De Pauw, & Kloots (2004) examined regional dialects of Dutch. Their study divided 160 speakers into eight dialectal groups, four dialects of Dutch as spoken in the Netherlands, and four other Dutch dialects from Flanders (in Belgium.) Fifteen-minute interviews with each speaker were analyzed for both speaking and articulation rate (which was called speech rate overall.) This region was interesting because of the pluricentric nature of the Dutch language – the language was similar enough in both countries in terms of structure and codified norms, but different enough to be certain that the speakers do not share all of the same "norms and linguistic variables" (p. 298). This eliminated the problem of linguistic backgrounds being simply too similar

to adequately compare them, but also the problem of different languages being too diverse in factors such as structure and word length to know if the speakers are just constrained by their language in terms of how fast they talk.

Results of this study demonstrated that speech rate did not change significantly among regions (only one significant difference was found in Netherlands, being that the speakers from the linguistic center of Randstad had a higher articulation rate than the others.) A significant difference was found between the countries, with all regions of the Netherlands having a faster speech rate than those in Belgium. The possible reason for this difference was the frequency of usage of the language. Standard Dutch is the main language in the Netherlands, while most of the Dutch speakers in Belgium used another dialect or variety of the language, that is quite different from the standard, just as often, or more often. This could cause a slower speaking rate just because they are less fluent or comfortable with the dialect. Other factors such as the speakers' possibly differing views on the level of formality of the interviews are considered, but Verhoeven, De Pauw, & Kloots do not rule out that over time the dialect of Dutch in the Belgium regions has become intrinsically slower and that norm has been passed down generations.

Age as a factor in speaking rate is more established than the other factors of dialect and gender. Multiple studies have shown the tendency for older speakers to talk more slowly than younger adults. Ramig (1983) considered physiological condition as a factor of age in measuring speaking rate. Physiological condition of the 48 speakers in three age groups (25-35, 45-55, and 65-75 years old) was determined by measuring "resting heart rate, resting systolic and diastolic blood pressure, percent fat and forced vital capacity" (p. 220). "Good" and "poor" physiological condition groups were formed

within each age group. Reading and speaking rates were then calculated (from the data collected when speakers read the "Rainbow Passage" and then spoke spontaneously for thirty seconds) and the results showed that there was an impact of physiological condition on speaking rate. For both reading and speaking rate, it was found that the rate of the older speakers was significantly slower than that of the younger group (the middle group did not play a significant factor in comparison with either group.) These differences were also more remarkable when comparing only those grouped in "poor" physiological condition. For example, young speakers in "good" condition had a speaking rate of 3.62 syllables per second, while older speakers in "good" condition had a rate of 3.28 syll/second. On the other hand, the young speakers in "poor" condition had a speaking rate of 3.89 syllables per second while the older speakers in "poor" condition's rate was 2.64 syll/second, a disparity of a much larger margin. Ramig cited physiological factors such as "visual acuity, processing time, general neuromuscular slowing, peripheral degeneration of the speech mechanism, and psychosocial variables" (p. 224) as possible reasons why the physical condition of an elderly person affected their speaking rate. An additional possible factor is differing levels of pulmonary functioning of the young speaker and elderly speakers.

Smith, Wasowicz, & Preston (1987) also looked at age as a factor of speaking rate. Their study separated the speakers by age (ten speakers 24-27 years old, ten speakers 66-75 years old) and did not separate on the basis of physiological condition. The health of the elderly speakers was discussed however, and it was noted that although they were "healthy, active, and living in their own homes" (p. 523), that some had medical histories, such as glaucoma, myocardial infarction, use of a pacemaker, and

operations for various cancers. They concluded that these speakers represented "the upper, middle portion of a large range of possible health conditions among the elderly" (p. 523). In several different sentence and word repetition tasks, about 200 utterances were produced by each speaker. Sentence durations were found to be about 22% longer for older adults than young adults. When speakers were asked to repeat sentences at a fast rate, the older adults' durations were on average 26% longer. However, the fast rate of the older adults was 13% faster than the normal rate of the young adults, showing that older adults *can* produce faster speech than normal adults, but do not do so when speaking normally. These results showed that older adults do tend to have a slower speaking rate than young adults, but could not address the exact causes for why this would be.

1.6. Anticipatory Shortening

Recently, the affect of "anticipatory shortening" was studied in depth as a function of speaking rate by Quené (2008). This proposal states that if a speaker anticipates having to say a lot of syllables in a run he will therefore decrease the length of each syllable. This will in turn increase speaking rate. The study also looked at regional factors, and used the same corpus material as the preceding study, Verhoeven, De Pauw, & Kloots (2004). Interviews conducted with each of the 160 speakers were analyzed by phrase for with-in and between-speaker factors. The data showed that speakers from Flanders (Belgium) had a slower speaking rate than those from The Netherlands, and the four dialects within each area differed as well. However, results showed far greater within-speaker variance than between-speaker differences. Particularly, the speakers

from Flanders varied their tempo away from the average more so than The Netherlands speakers. Furthermore, speakers from The Netherlands used shorter phrases, as the elderly had a tendency to do as well. Quené hypothesized that phrase length is important in measuring speaking rate, but a person's average phrase length could be influenced by age, "and perhaps also by other between-speaker predictors: sex, country, and region" (p. 1109). This study is important to consider, but for the purposes of the present study, phrase length as a predictor of speaking rate will not be the focus.

1.7. Variation in Vowel Duration

Jacewicz, Salmons, & Fox's 2007 study on vowel duration of Wisconsin, North Carolina, and central Ohio speakers may provide some insight to predict the speaking rate of these two regions. In the study, 54 speakers (ages 20-34, eighteen speakers from each region) were asked to read a total of ninety sentence pairs, and to place emphasis on the designated target word. Vowels in three categories of emphasis (high, intermediate, or low) were analyzed for within- and between-speaker factors. It was found that there was a significant difference between the average length of vowels depending on level of emphasis, but the study also showed that, no matter the level of emphasis, "North Carolina vowels were significantly longer than either Ohio or Wisconsin vowels" (p. 373). By gender, vowels spoken by females were longer than those by males in North Carolina, but in Wisconsin were found to be either shorter or equal in duration. Overall, the effects of gender were not as strong as those of dialect, and concluded to seen as a "tendency rather than a true effect" (p. 378). Within dialects, a significant effect of consonantal context was also observed. Both dialects and genders showed greater vowel

length (for all vowels) when they occurred before voiced consonants (rather than voiceless.) In noting that the vowel lengths differ among within-speaker and between-speaker factors, this leads to the possibility of this variation in vowel length to affect the overall speaking rate of these speakers.

Chapter 2

Methodology

Two separate sets of data were analyzed for the purposes of the study, a set of spontaneous talks and a set of read sentences. These two sets served as a testing material for investigation of spontaneous speaking rate and reading rate. This data was previously collected from subjects at the University of Wisconsin-Madison and Western Carolina University for the larger project "Cross-generational sound change in America English" conducted at the Speech Perception and Acoustics Labs at The Ohio State University.

2.1. Speakers

The speakers fell into two age categories (a younger group, ages 20-34 and an older group, ages 51-65.) All of the Wisconsin speakers were born and raised in the Madison area - which consists of Dodge and Dane counties (south-central Wisconsin), and in North Carolina, speakers were born and raised in the Sylva, Cullowhee, and Waynesville areas – Jackson, Swain, and Haywood counties (western North Carolina.) The speakers from Wisconsin reported living in a more suburban location, and the North Carolina speakers reported living in a more rural environment and small towns. All of the young speakers were students at either the University of Wisconsin-Madison or Western Carolina University. Older speakers came from a variety of educational backgrounds, including high school and college-level attainment.

For sentences, there were a total of 76 speakers, 38 from Wisconsin and 38 from North Carolina. Each dialect group was separated by gender, with even numbers of females and males (19 males and 19 females from each region). Lastly, groups were

separated by age, with nine in each group falling into the 20-34 year old range, and ten in the 51-65 year old range. Therefore, overall the groups were: young (20-34) Wisconsin females, older (51-65) Wisconsin females, young Wisconsin males, older Wisconsin males, young North Carolina females, older North Carolina females, young North Carolina males, and older North Carolina males.

Eighty speakers participated in the interview portion of the study. Fifty-seven of these speakers also participated in reading the sentences. All of the older speakers participated in both sets of the study, while almost all of the young North Carolina speakers were different for the interviews, as well as some of the young Wisconsin speakers. The eighty speakers were divided into the same groups as for the sentences, however in the four young adult groups there were ten subjects instead of nine.

2.2. Stimuli

For the sentence portion, each speaker was asked to read two sets of sixty sentence pairs. In each sentence there was a "target word" that contained a particular vowel of interest. The main sentence stress varied in order to create different levels of emphasis (high, intermediate, or low) on the target word. The speaker was asked to put stress on the capitalized word. In the study "Vowel Duration in three American English dialects" (Jacewicz, Salmons, Fox 2007) these systematic stress changes were used to examine the contextual effects on vowel duration in respect to dialect and gender. An example of a set of sentence pairs with the target word "bits" is shown below (the task in its entirety can be found in the appendix):

JANE knows the small bits are sharp. No! JOHN knows the small bits are sharp.

John FEELS the small bits are sharp. No! John KNOWS the small bits are sharp.

John knows the SOFT bits are sharp. No! John knows the SMALL bits are sharp.

John knows the small SCREWS are sharp. No! John knows the small BITS are sharp.

John knows the small bits are DULL. No! John knows the small bits are SHARP.

Spontaneous talks consisted of the speaker answering general questions posed by an interviewer about his/her life. Speakers often talked about their family, their family history, their job, schooling, and hobbies. Some told stories, particularly in regard to their dialect, and experiences they have had because of it. The interviews lasted anywhere from approximately two minutes to almost 20 minutes in length. In general, the North Carolina interviews lasted longer (with most averaging between 10-15 minutes) and the Wisconsin interviews were shorter (most between 5-10 minutes.) Due to time constraints, only the first five minutes of each interview was analyzed, an amount of time that was considered sufficient to obtain a measurement of speaking rate.

2.3. Recording Procedure

Recording of sentences was conducted using a custom program that was written in Matlab. The speaker was asked to read the sentence pair that appeared on a computer monitor, which were presented in random order. The speaker spoke to a head-mounted microphone (Shure SM10A) that was located 1 inch from their lips. Speakers were asked to repeat the sentence pairs if they made an error, such as mispronunciation, disfluency, or if they did not correctly stress the target word (by either stressing a different word in the sentence or simply a lack of stress at all.) The stimuli were recorded onto a hard drive disk at a 44.1-kHz sampling rate.

As for the interview portion, speakers were recorded using the Adobe Audition speech analysis program.

2.4. Acoustic Measurements

Adobe Audition was used to listen to the interviews and sentences. For sentences, Adobe Audition was used to measure the second sentence *only* of each of the 120 sentence pairs, also excluding the "No" at the beginning of each second sentence. The onsets and offsets of each sentence were entered in an excel file, as well as the number of syllables in the same file (each sentence contained seven syllables.) Sentences were double checked for speaker errors. Errors were few because of repeating a sentence pair to get the more fluent reading, however when one was found, its duration was added to the excel file containing onsets and offsets of the whole sentence and number of syllables. Before reading rate was calculated (sentence duration divided by number of syllables) the pauses were taken into account by subtracting the duration time of the pause from the duration time of the full sentence.

All of the interviews had been previously transcribed and the first five minutes of the transcriptions were broken down into phrases. Isolating phrases of fluent speech was the first step in determining articulation rate. Since articulation rate is a measure of the pace at which speech segments are produced, no pauses should be included. Therefore for the present purpose, a phrase was defined as a run of speech consisting of five or more syllables that was produced without pause or hesitation. For example, if a speaker interjected a run with a drawn-out filler such as "umm" or "uh", or if he/she took a breath, the run would be counted as two different phrases. If a filler was fluent, however,

it could remain in the phrase. Each phrase was numbered, indented, and colored so as to stand apart from the utterances that were not phrases. The total number of phrases per speaker usually fell between forty and eighty, with a majority of them in the 60-75 range. A few examples of how these transcriptions were arranged are shown below:

older North Carolina female speaker:

52) and to tell a funny story about,
uh, my speech,
we have a,
53) a mountain pasture where we keep our cows, our cattle in the uh,
spring and summer,
and this,
uh,
54) man from Florida came up and bought the place next to it,
55) and so, we were up there checking on the cattle one day and he,
56) we stopped and started talking to him and,
57) he got acquainted with us and uh,
58) he was just delighted with my speech and I didn't think there was anything wrong with it but,
59) I did not get the feeling that he was making fun of me and,
60) he kept saying that the next time they came up from Florida that he was going to,
um,
have,
61) bring his wife and let her hear me talk,

older North Carolina female speaker:

Well I guess that
34) growing up with my grandparents was,
uh,
one of the,
35) treasures that I have looking back on my life, another treasure I have was that,
36) as I got older, we had this older man and lady that lived in our farm house and helped take care of the
37) cattle and work on the farm…
38) and he was the type of person that,
39) in today's world, could of made a fortune writing books because he was forever telling some kind of a story,
40) and what's so interesting about it is he told the same story no matter how many times he told it,
41) a year apart and the story would be exactly the same

older Wisconsin male speaker:

27) coming from a small town Milwaukee was a little more interesting. I actually went there
28) just to finish school but I stayed,
got a job um,
29) I had been working,
uh,
30) as a janitor for most of the time I was there,
um
31) getting around town, and I, I took the bus to get wherever I needed to go which really helped a lot.
I learned
32) the city of Milwaukee real quick.
And um,
33) had a few friends there so it was pretty comfortable.

Once the phrase transcriptions were complete, an excel file was created for each speaker. In separate columns, the number of syllables in each phrase was entered, as well as the onset and offset time of the phrase (the second/ millisecond mark the phrase occurred in the interview.) The onset was subtracted from offset to calculate phrase duration, then the number of syllables in a phrase was divided by the phrase duration to attain the articulation rate in syllables per second of every phrase spoken by every speaker.

2.5. Statistical Analysis

Two separate univariate analyses of variance (ANOVA) were used to initially assess the results. In the first ANOVA, the dependent variable was articulation rate in reading, with between-subject factors of dialect, age, and gender. The second ANOVA used articulation rate in talks as the dependent variable with the same between-subject factors as in the first.

Chapter 3

Results and Discussion

3.1. Read Versus Spontaneous Speech

The overall mean articulation rate in read sentences was 3.40 syll/s whereas the rate in spontaneous talks was 5.12 syll/s. This equates to a 66.4% difference between the read and spoken productions, large enough to indicate that in general, articulation rate in reading is much slower than in free speech. These results support those of earlier studies. The main interest of this study, however, was to determine whether there were significant differences within each type of production as a function of speaker dialect, age, and gender. The following analyses were based on the ANOVA assessments previously described, and any other additional analyses conducted are detailed below.

3.2. The Effects of Speaker Dialect

The results show that speaker dialect had a significant effect on both read sentences and spontaneous speech rates, with Wisconsin speakers demonstrating the faster articulation rate in each as compared with North Carolina speakers. Figures 1 and 2 display these differences as a function of speaker dialect.

Specifically, for the articulation rate in read sentences, the overall means were 3.54 syll/s for Wisconsin speakers and 3.26 syll/s for North Carolina, meaning Wisconsin speakers read at a rate 8% faster than North Carolina. The main effect of dialect was significant [$F(1, 68) = 12.8$, $p = 0.001$, $\eta^2 = 0.158$].

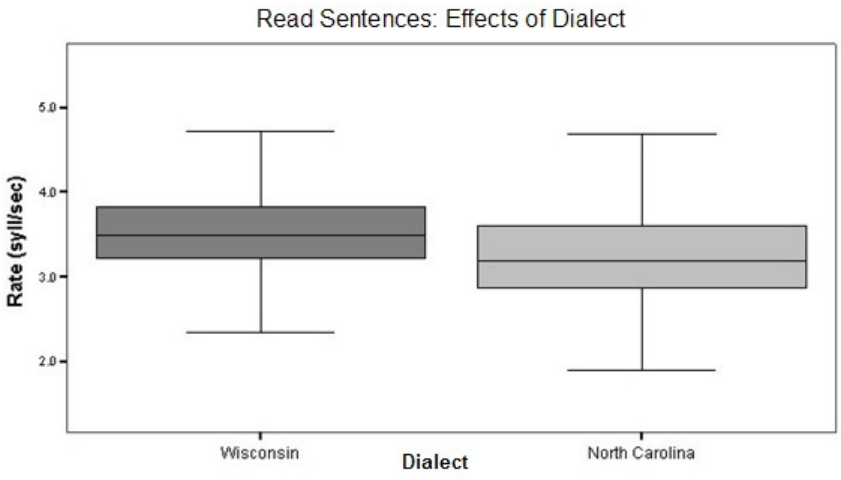

Figure 1: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate in read sentences for Wisconsin and North Carolina speakers.

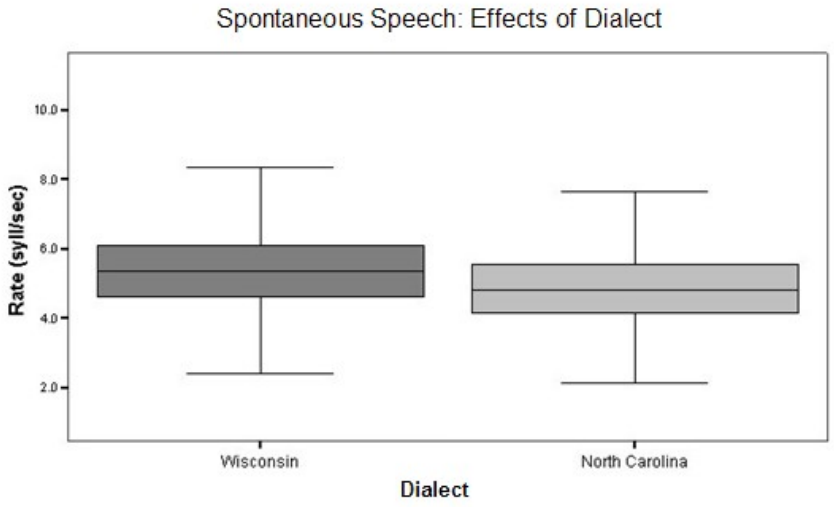

Figure 2: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate in spontaneous speech for Wisconsin and North Carolina speakers.

For spontaneous speech, the second ANOVA showed a significant main effect of dialect [$F(1, 72) = 28.5$, $p < 0.001$, $\eta^2 = 0.284$], with overall mean rates of 5.41 syll/s for Wisconsin and 4.89 syll/s for North Carolina speakers. This is a 9.6% difference in

articulation rate between the two. These findings clearly show that the articulation rate, whether in reading or speaking, is higher for the regional variety of American English spoken in Wisconsin as compared to North Carolina.

3.3. The Effects of Speaker Age

Figures 3 and 4 indicate the mean articulation rate for young and older adults in Wisconsin and North Carolina for read sentences and spontaneous speech rates. It can be observed that young adults tend to speak faster than older adults in both types of productions in Wisconsin. However, North Carolina young adults tend to speak faster only in read sentences but not in spontaneous talks.

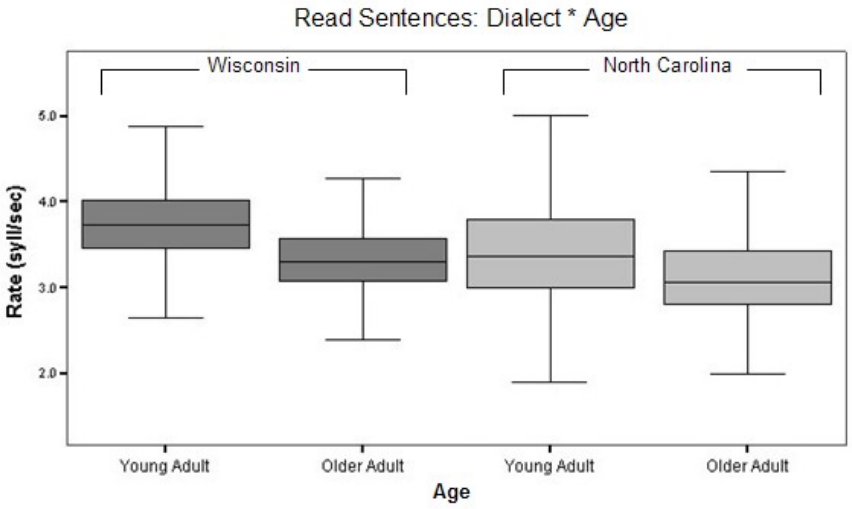

Figure 3: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate in read sentences for older and younger Wisconsin and North Carolina speakers.

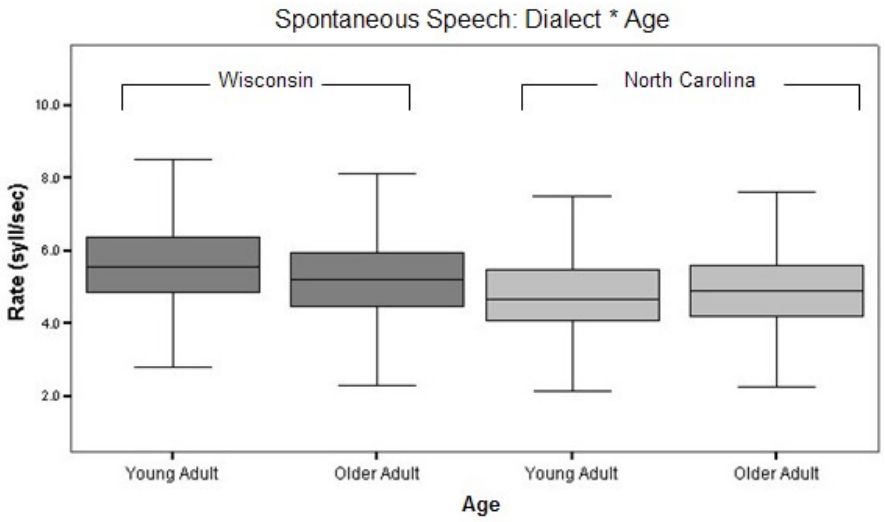

Figure 4: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate in read sentences for older and younger Wisconsin and North Carolina speakers.

The ANOVA results for read sentences showed a significant main effect of age [$F(1, 68) = 21.0$, $p < 0.001$, $\eta^2 = 0.236$], indicating that young adults' rate in read sentences was 10% faster than that of older adults (3.58 syll/s vs. 3.23 syll/s). For spontaneous talks, however, the main effect of age was not significant, indicating no differences in articulation rate between the young and older adults. Because the results for Wisconsin suggested some differences between the two groups (see Figure 4), we conducted additional two-way ANOVAs with the between-factors age and gender separately for Wisconsin and North Carolina talks. The results for Wisconsin showed a significant main effect of age [$F(1, 36) = 5.26$, $p = 0.028$, $\eta^2 = 0.128$] although the effect size was small. Young Wisconsin adults were shown to speak 6.4% faster than older adults (5.60 syll/s vs. 5.24 syll/s). For North Carolina, the main effect of age was not significant.

Overall, the results show that the articulation rate for read sentences is faster for young speakers than older speakers across both dialects. The articulation rate for spontaneous speech, however, is not as consistent. Results showed that the young adults tend to speak faster than older adults in Wisconsin, but not in North Carolina, where the young and older adults do not show differences in the articulation rate (4.79 and 4.83 syll/s, respectively.)

3.4. The Effects of Speaker Gender

Differences in articulation rate as a function of gender were very small, as is shown in Figures 5 and 6. As a general tendency, men spoke faster than women both in read sentences (3.47 syll/s vs. 3.33 syll/s) and in spontaneous talks (5.23 syll/s vs. 5.04 syll/s) but the differences were very modest. For read sentences, the articulation rate for men was 4% faster and for spontaneous talks, it was 3.6%. The results of the two ANOVAs for either articulation rate in read sentence or spontaneous speech showed no significant main effect of speaker gender.

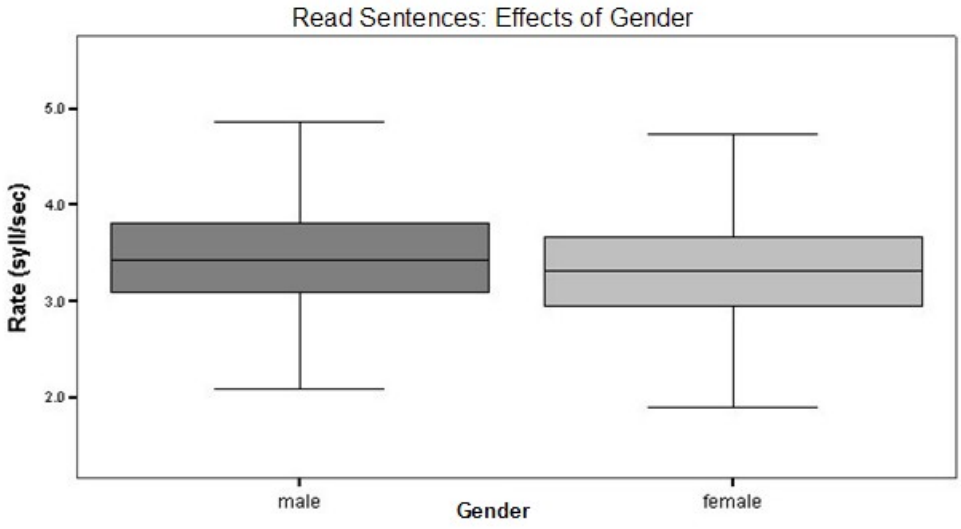

Figure 5: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate in read sentences for gender.

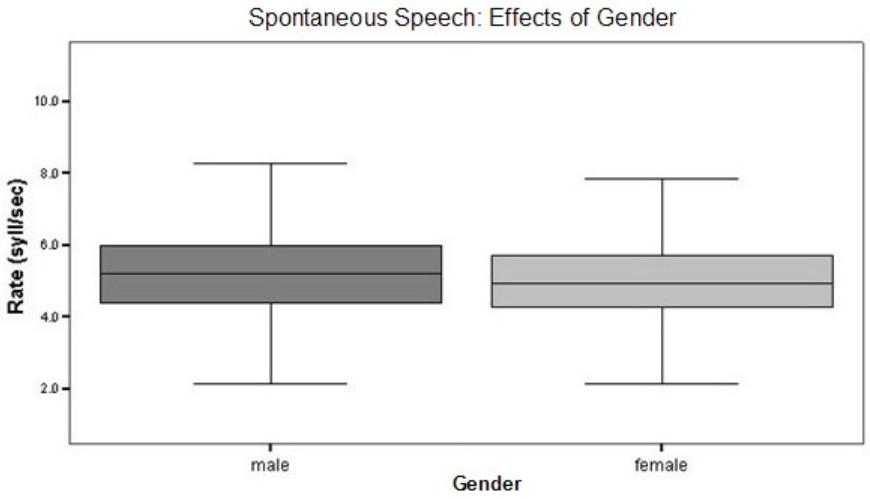

Figure 6: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate in spontaneous speech for gender.

Two-way ANOVAs were also conducted for each dialect with the between-subject factors gender and age. The main effect of gender on either articulation rate (reading or spontaneous talks) was again not significant for either Wisconsin or North Carolina (see Figures 7 and 8 below.)

Figure 7: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate in read sentences for dialect and gender. While dialect was found to be a significant factor, gender was not.

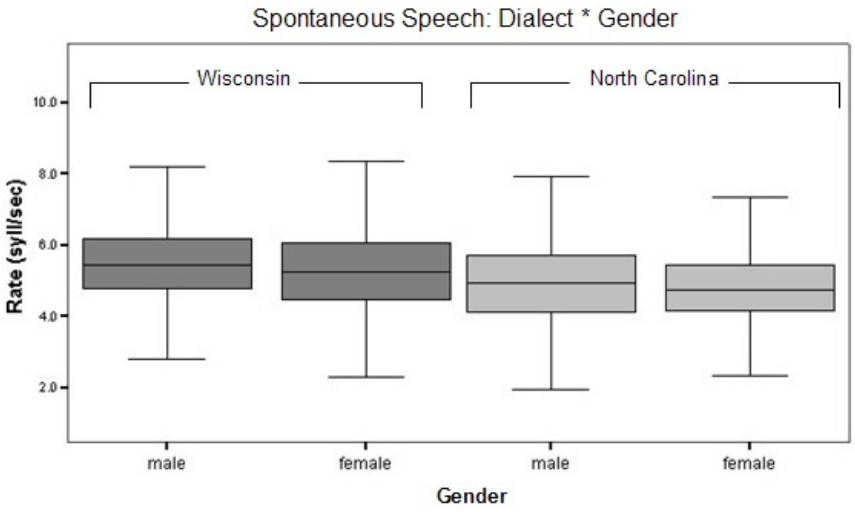

Figure 8: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate in spontaneous speech for dialect and gender. As with read sentences, gender was not a significant effect.

However, there was one significant age by gender interaction for read sentences for North Carolina although its effect size was small [$F(1, 34) = 4.56$, $p = 0.040$, $\eta^2 = 0.118$]. This interaction indicated greater differences in the articulation rate between young men and young women as compared to older men and older women, as can be seen in Figure 10. This interaction did not occur in Wisconsin young men and women's rates, which can be seen in Figure 9. A subsequent 2-tailed independent samples t-test showed that the difference for young North Carolina adults was significant [$t = 2.7$, $p = 0.015$], and showed that young men in North Carolina read 14% faster than young women (3.67 syll/s vs. 3.14 syll/s). The difference in articulation rate for read sentences between older North Carolina men and older women was not significant (3.12 and 3.15 syll/s, respectively).

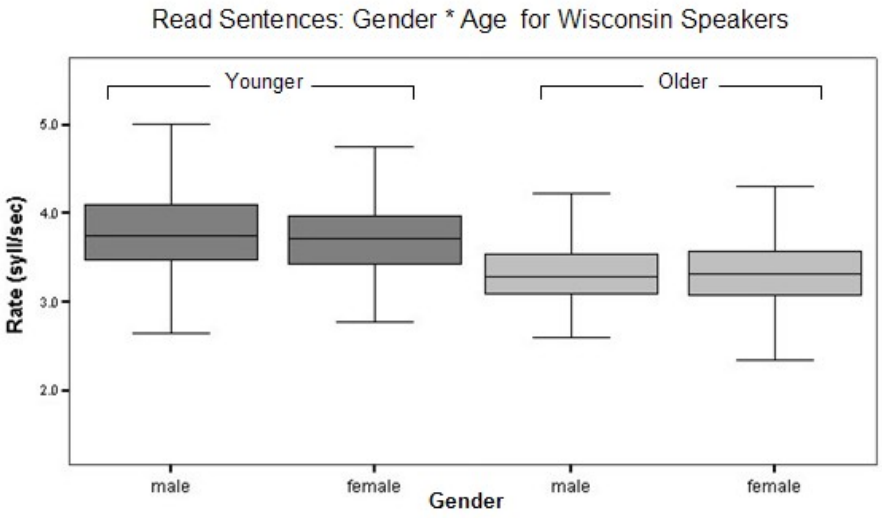

Figure 9: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate read sentences by gender and age for Wisconsin speakers. There were no significant results for this interaction.

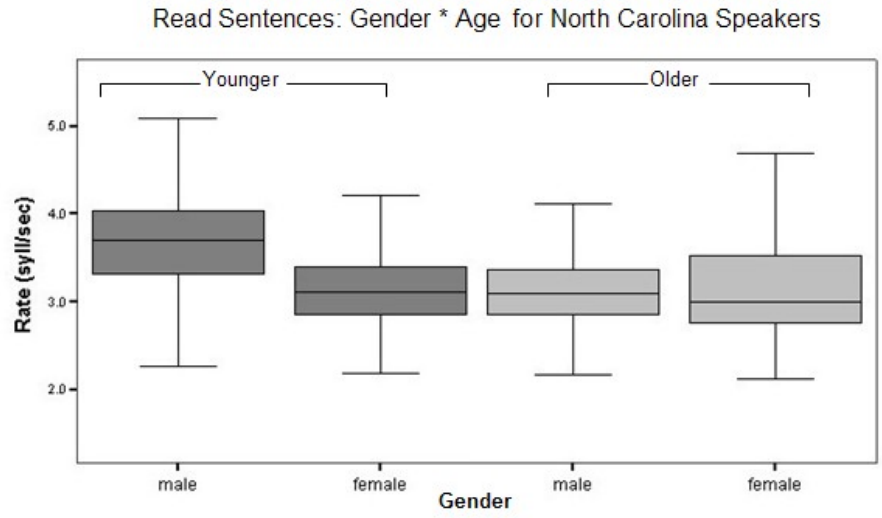

Figure 10: A Box-and-Whisker Plot showing the central tendency and dispersion of articulation rate read sentences by gender and age for North Carolina speakers. Young North Carolina speakers showed an age by gender interaction. The young male North Carolina speakers read significantly faster than young female North Carolina speakers.

Overall, the effects of speaker gender on articulation rate were too small to reach statistical significance. Generally, men tended to speak modestly faster than women with

one exception in North Carolina, where young men read significantly faster as compared to young women.

3.5. Discussion

This study examined the verity of the stereotype that southern speakers talk more slowly than northern speakers by examining the speaking rates of spontaneous speech and read sentences of two large groups of speakers, one group from south-central Wisconsin and the other from western North Carolina. Additional goals of the study were to make cross-generational and cross-gender comparisons of speaking rate.

Findings in this study show that the measurement of articulation rate is a significant factor in varying speaking rates. The present results support those of Miller, Grosjean, and Lomanto (1984) who first showed changeability in articulation rate. Present results also correspond with the results of Jacewicz, Salmons, and Fox (2007) who found that vowels are significantly longer in North Carolina than in Wisconsin. It is reasonable to assume that the difference in lengths of vowels contributes to the difference in articulation rates found in the present study.

Results showed that there was a significant difference between the articulation rate for read sentences and spontaneous speech for all speakers, with the rate of reading being much slower, a result that was not unexpected. These findings support those of previous studies where it was found that read speech was slower than spontaneous speech due to increased reduction in spontaneous speech (Crystal & House 1982) and prosodic differences (Hirose & Kawanami 2002). However it is inconsistent with the results of Howell & Kadi-Hanifi (1991) who although they found differences in reading and

speaking rate, their study showed speakers reading more quickly than speaking spontaneously. This may be because pauses were included in this study, and perhaps people were pausing to remember details about the room they were asked to describe, where a detailed memory would not be a factor while reading. It is also important to note that with these three studies, the same speakers were used for all sets of material, as with the majority of the speakers in the present study (the present study however, had analyzed a much greater number of speakers than the other studies, which had fourteen, one, and six speakers respectively.)

Dialect was also significant for each of these rates as well, with Wisconsin speakers having a significantly faster articulation rate than North Carolina speakers for both spontaneous talks and read sentences. The results are not consistent with those of Ray & Zahn (1990) who found that region had little effect on speaking rate. This could be because of the differences in regions represented. Ray and Zahn (1990) measured the speaking rates of students attending universities in the Pacific Northwest, the Southwest, and the Upper Midwest. Perhaps these regions do not differ in speaking rate as greatly as the Inland North region of Wisconsin and the Appalachian region in Western North Carolina. Or perhaps the large two-three state area the regions covered in the 1990 study made results more general than the three-to-four county area that was studied presently. Lastly, Ray and Zahn made their measurements in words per minute, while the present studied looked at syllables per second – perhaps the more detailed measurements made here enabled differences to become more visible.

The present results also supported Byrd (1994) who found that out of eight dialect regions in TIMIT database, the speakers from the South had the slowest speaking rate.

Though Byrd found the results to be inconclusive due to the broad dialect regions represented, again, the results of the present study were derived from data from very specific dialect regions, so the presence of a slower speaking rate among southern speakers can be drawn more conclusively here.

Age also proved to be a significant factor, at least with Wisconsin speakers. Young Wisconsin speakers spoke significantly faster than older Wisconsin speakers for both spontaneous and read sentence rates, but in North Carolina younger speakers had faster rates only for read sentences, and the differences for spontaneous speaking rates were not significant. The results for Wisconsin support the findings of Ramig (1983) who found that for both read sentences and spontaneous speech rates, older speakers were slower than younger speakers. The results of North Carolina, however, do not support Ramig's study because younger speakers only read faster than older speakers and there were no differences in the spontaneous speaking rate. Perhaps the slower rate of sentence reading could be related to the educational levels of the older versus younger North Carolina speakers (all of the younger speakers were college-educated and the levels of education varied among the older speakers) and have less to do with an inherently different rate in speaking.

Results for gender were more mixed than for the other two factors, with very small differences shown. Overall, men tended to speak faster than women but the only significant result was between young North Carolina men and women, with the young men reading faster than the young women. That one significant result agrees with Byrd (1994) who found that men speak significantly faster than women (6.2%), but overall this study did not find significant results when comparing gender. As Byrd identified, the

weakness in the database she used, TIMIT, was that the group of speakers consisted of 70% male, meaning that the women were underrepresented. Also, the speakers in the database were only reading sentences, not speaking spontaneously. Lastly, pauses were not edited out of her data. The present study used equal numbers of male and female speakers, and compared two types of production, editing out pauses to obtain articulation rate, and therefore overcome the weaknesses that were present in Byrd (1994).

Of the reported percentages of the differences between speaking rates, most are above the previously mentioned Just Noticeable Difference for speaking rate (which was found to be around 5%). For example, the difference between the dialects was 8% for read sentence rates, and 9.6% for spontaneous speech rates, while the difference between young and older speakers was 10% for read sentence rates, and 6.4% for young and older Wisconsin speakers for articulation rate in spontaneous speech. For gender, however, the differences between men and women speakers fell below 5% (4.6% for read sentences and 3.6% for spontaneous speech overall.) There was one exception for gender - the difference between young men and women in North Carolina was 14% for read sentences. Therefore, we hypothesize that listeners likely could hear the differences in speaking rates between the dialect and age groups, but probably could not in the gender group overall.

The strengths of this study lie in the large amount of data that was collected from the large numbers of speakers from two specific regions. This allows for the most detailed dialect study on speaking rate to date. Also, because a majority of the same speakers participated in the sentence reading portion as well as the spontaneous speaking section, the results allowed for comparison of the same person to see the effect of the

type of production on the speaking rate. Overall, the results of this study were very consistent. Less consistent results, such as the differences in young men and women speakers in North Carolina but no differences in Wisconsin, and the lack of difference in spontaneous speaking rate in older and young North Carolina speakers, while there was as significant difference in Wisconsin speakers could benefit from further study, perhaps by including the Ohio speakers from the larger study in the data.

Second, this study did not take into account the phrase lengths used by individual speakers to determine if anticipatory shortening was a factor in the differing speaking rates, as was discussed in the Verhoeven, De Pauw, & Kloots (2004) study. This aspect could be investigated further in another study with the present data to see if it is a more important effect than dialect, age, or gender.

Overall this study is important in the field because it is the first to compare reading and speaking rate for these two dialects, and to include gender and age. The results were reliable and should be considered when conducting studies of this kind in the future.

Chapter 4

Acknowledgments

I would like to thank Dr. Fox and Dr. Jacewicz for all of the time, guidance, and knowledge they provided in order to support me in completing this project. I would also like to thank Dr. Weisenberger for introducing me to the opportunity to research. Thank you to everyone working in or with Speech Perception and Acoustics Labs, for the work and support that allowed me to finish on time! In particular, I would like to thank my fellow undergraduate researcher Samantha Lyle for her support and friendship throughout this year. Lastly, thank you to Dr. Grinstead, for serving on my Defense Committee.

This project was supported by the Ohio State University Colleges of Arts and Sciences and College of Social and Behavioral Sciences.

Chapter 5

References

Byrd, D. (1994). Relations of Sex and Dialect to Reduction. *Speech Communication*, *15*, 39-54. Retrieved February 26, 2008, from http://www-rcf.usc.edu/~dbyrd/byrd_speechComm15_1994.pdf

Cran, W., & MacNeil, R. (2005). *Do You Speak American?* Retrieved April 16, 2008, from http://www.pbs.org/speak/education/curriculum/college/regional/

Crystal, T. H., & House, A. S. (1982). Segmental Durations in Connected Speech Signals: Preliminary Results. *J. Acoust. Soc. Am.*, *72*, 705-716.

Goldman-Eisler, F. (1961). The Significance of Changes in the Rate of Articulation. *Language and Speech*, *41*, 171-188.

Hewlett, N., & Rendall, M. (1998). Rural versus Urban Accent as an Influence on the Rate of Speech. *Journal of the International Phonetic Association*, *28*, 63-71. Retrieved September 17, 2007, from CSA Linguistics and Language Behavior Abstracts database.

Hirose, K., & Kawanami, H. (2002). Temporal Rate of Change of Dialogue Speech in Prosodic Units as Compared to Read Speech. *Speech Communication*, *36*, 97-111.

Howell, P., & Kadi-Hanifi, K. (1991). Comparison of Prosodic Properties Between Read and Spontaneous Speech Material. *Speech Communication*, *10*, 163-169.

Jacewicz, E., Salmons, J., & Fox, R. A. (2007). Vowel Duration in Three American English Dialects. *American Speech*, *82*, 367-385.

Miller, J. L., Grosjean, F., & Lomanto, C. (1984). Articulation Rate and Its Variability in Spontaneous Speech: A Reanalysis and Some Implications. *Phonetica*, *41*, 215-225. Retrieved August 26, 2007, from NCBI database.

O'Neil, D. (2006, August 31). *What is Language?* Retrieved April 16, 2008, from http://anthro.palomar.edu/language/language_2.htm

Quené, H. (2007). On the Just Noticeable Difference for Tempo in Speech. *Journal of Phonetics*, *35*, 353-362.

Quené, H. (2008). Multilevel Modeling of Between-Speaker and Within-Speaker Variation in Spontaneous Speech Tempo. *J. Acoust. Soc. Am.*, *123*, 1104-1113.

Ramig, L. A. (1983). Effects of Physiological Aging on Speaking and Reading Rates. *Journal of Communication Disorders*, *16*, 217-226.

Ray, G. B., & Zahn, C. J. (1990). Regional Speech Rates in the United States: A Preliminary Analysis. *Communication Research Reports*, *7*, 34-37.

Smith, B.L., Wasowicz, J., & Preston,J. (1987). "Temporal characteristics of the speech of normal elderly adults." *Journal of Speech and Hearing Research*, *30,* 522-529.

Tsao, Y., Weismer, G., & Iqbal, K. (2006). Interspeaker Variation in Habitual Speaking Rate: Additional Evidence. *Journal of Speech, Language, and Hearing Research, 49,* 1156-1164.

Verhoeven, J., De Pauw, G., & Kloots, H. (2004). "Speech Rate in a Pluricentric Language: A Comparison Between Dutch in Belgium and the Netherlands." *Language and Speech, 47,* 297-308.

Whiteside, S. P. (1996). Temporal-Based Acoustic-Phonetic Patterns in Read Speech: Some Evidence for Speaker Sex Differences. *Journal of the International Phonetic Association*, *26*, 23-40.

Index of Figures

Appendix

The following sets of sentences were recorded by each speaker. All 2-set sentences were randomly presented to the subject in two stimulus lists. The sentences in which the main stressed falls on the first and second word position, respectively, served as distracters and were not included in the final analyses.

Vowels before a voiceless consonant in a word

bits

JANE knows the small bits are sharp.
No! JOHN knows the small bits are sharp.

John FEELS the small bits are sharp.
No! John KNOWS the small bits are sharp.

John knows the SOFT bits are sharp.
No! John knows the SMALL bits are sharp.

John knows the small SCREWS are sharp.
No! John knows the small BITS are sharp.

John knows the small bits are DULL.
No! John knows the small bits are SHARP.

baits

MOM said the dull baits are best.
No! DAD said the dull baits are best.

Dad THINKS the dull baits are best.
No! Dad SAID the dull baits are best.

Dad said the BRIGHT baits are best.
No! Dad said the DULL baits are best.

Dad said the dull HOOKS are best.
No! Dad said the dull BAITS are best.

Dad said the dull baits are WORST.
No! Dad said the dull baits are BEST.

bets

FRANK said the small bets are low.
No! JOHN said the small bets are low.

John THOUGHT the small bets are low.
No! John SAID the small bets are low.

John said the BIG bets are low.
No! John said the SMALL bets are low.

John said the small POTS are low.
No! John said the small BETS are low.

John said the small bets are HIGH.
No! John said the small bets are LOW.

bats

SAM said the small bats are fast.
No! DOC said the small bats are fast.

Doc THINKS the small bats are fast.
No! Doc SAID the small bats are fast.

Doc said the LARGE bats are fast.
No! Doc said the SMALL bats are fast.

Doc said the small BIRDS are fast.
No! Doc said the small BATS are fast.

Doc said the small bats are SLOW.
No! Doc said the small bats are FAST.

bites

JANE thinks the small bites are deep.
No! SUE thinks the small bites are deep.

Sue KNOWS the small bites are deep.
No! Sue THINKS the small bites are deep.

Sue thinks the LARGE bites are deep.
No! Sue thinks the SMALL bites are deep.

Sue thinks the small CUTS are deep.
No! Sue thinks the small BITES are deep.

Sue thinks the small bites are WIDE.
No! Sue thinks the small bites are DEEP.

Vowels before a voiced consonant in a word

bids

BOB thinks the fall bids are low.
No! TED thinks the fall bids are low.

Ted KNOWS the fall bids are low.
No! Ted THINKS the fall bids are low.

Ted thinks the SPRING bids are low.
No! Ted thinks the FALL bids are low.

Ted thinks the fall SALES are low.
No! Ted thinks the fall BIDS are low.

Ted thinks the fall bids are HIGH.
No! Ted thinks the fall bids are LOW.

bades

(The nonsense word *bade* was explained to the speaker as indicating "a brand of knife, a brand name.")

TOM says the dull bades are cheap.
No! TED says the dull bades are cheap.

Ted THINKS the dull bades are cheap.
No! Ted SAYS the dull bades are cheap.

Ted says the SHARP bades are cheap.
No! Ted says the DULL bades are cheap.

Ted says the dull FORKS are cheap.
No! Ted says the dull BADES are cheap.

Ted says the dull bades are WEAK.
No! Ted says the dull bades are CHEAP.

beds

TOM said the tall beds are warm.
No! ROB said the tall beds are warm.

Rob THINKS the tall beds are warm.
No! Rob SAID the tall beds are warm.

Rob said the SHORT beds are warm.
No! Rob said the TALL beds are warm.

Rob said the tall CHAIRS are warm.
No! Rob said the tall BEDS are warm.

Rob said the tall beds are COLD.
No! Rob said the tall beds are WARM.

bads

(The speaker was told that *bad* refers to "an error or mistake." For example, if someone makes an error, he or she might say "my bad" instead of "my mistake.").

NICK thinks the small bads are worse.
No! MIKE thinks the small bads are worse.

Mike KNOWS the small bads are worse.
No! Mike THINKS the small bads are worse.

Mike thinks the BIG bads are worse.
No! Mike thinks the SMALL bads are worse.

Mike thinks the small GOODS are worse.
No! Mike thinks the small BADS are worse.

Mike thinks the small bads are BEST.
No! Mike thinks the small bads are WORSE.

bides

(The nonsense word *bide* was explained to the speaker as indicating "a small animal, a type of dog.")

SUE thinks the small bides are cute.
No! JANE thinks the small bides are cute.

Jane KNOWS the small bides are cute.
No! Jane THINKS the small bides are cute.

Jane thinks the SHORT bides are cute.
No! Jane thinks the TALL bides are cute.

Jane thinks the small CATS are cute.
No! Jane thinks the small BIDES are cute.

Jane thinks the small bides are GROSS.
No! Jane thinks the small bides are CUTE.